Non-Asymptotic Bounds on the Performance of Dual Methods for Resource Allocation Problems

Simon Görtzen and Anke Schmeink, *Member, IEEE*

***Abstract*—In this paper, dual methods based on Lagrangian relaxation for multiuser multicarrier resource allocation problems are analyzed. Their application to non-convex resource allocation problems is based on results guaranteeing asymptotic optimality as the number of subcarriers tends to infinity. This work analyzes the workings and performance of dual methods for resource allocation problems with concave rate functions and a finite number of subcarriers. The core results are the convexity of resource allocation problems with subcarrier sharing and an upper bound on the number of subcarriers being shared. Based on these results, absolute and relative performance bounds are presented for dual methods when applied to the resource allocation problem without subcarrier sharing. The exemplary problems considered in this work are sum rate maximization with global and individual power budgets and sum power minimization with global and individual rate demands.**

***Index Terms*—Resource allocation, adaptive modulation, orthogonal frequency division multiple access (OFDMA), duality theory, convex optimization, combinatorial optimization.**

I. Introduction

MULTICARRIER communication systems are systems in which the available spectrum is separated into orthogonal communication channels, or subcarriers. A central entity is tasked to allocate the available subcarriers and power to users, which constitutes a mixed integer-continuous multicarrier resource allocation problem (RAP). This class of RAPs has been widely studied, with the assignment of non-interfering subcarriers in orthogonal frequency division multiple access (OFDMA) systems constituting a prominent practical example and research topic. Broadening the definition, any communication system that allows for the orthogonal division and assignment of resources across one or more dimensions encounters problems of resource allocation.

In a multicarrier system serving a single user, the optimal solution is given by classical bit-loading [1]. For the multiuser case, subcarriers have to be uniquely allocated to users, complicating the combinatorial optimization problem. The objective of this problem is to maximize an objective function which depends on the users' data rates and power consumption per subcarrier. This maximization problem has to be solved under one or more constraints regarding power consumption and/or data rate requirements. In [2], dual methods for non-convex multicarrier RAPs are introduced for concave as well as for non-differentiable, "discrete" rate functions. The foundation for these methods is given by previous advances in optimization theory [3], [4]. Since then, various authors have investigated different RAPs based on the arguments presented in [2]. While the main research focus is on concave rate functions [5]–[8], discrete utility functions are also analyzed [9]–[12], often in addition to the concave case. This list is not meant to be exhaustive but shows that dual methods are the state of the art to solve multicarrier RAPs. In [13], RAPs with discrete rate functions are analyzed and performance bounds for dual methods are provided for three exemplary problems. In addition to providing an absolute bound for RAPs with a finite number of subcarriers, these bounds formally prove that the relative performance reaches tends towards 100% for a growing number of subcarriers. This paper generalizes the results of [13] to multicarrier RAPs with concave utility functions. Consequently, the problems in this work are mixed integer-continuous rather than linear integer programs, which necessitates a broader range of mathematical tools.

Dual methods are a powerful tool to solve arbitrary RAPs. Their good performance is being attributed to the fact that the RAP satisfies a so-called time-sharing condition as the number of subcarriers N goes to infinity. As shown in [2], problems that satisfy this condition have zero duality gap. Thus, dual methods can be applied to obtain an asymptotically optimal solution to the original problem.

Practice has shown that dual methods perform very well for systems with a comparably small number of subcarriers. Exemplary numerical computations in [5] also support the idea that satisfactory performance of dual methods does not rely on large N. However, to the best of the authors' knowledge, the non-asymptotic case has not been formally analyzed regarding questions of duality gap, convexity, and optimality. The main purpose of this work is to offer a new perspective on dual methods for RAPs which allows for such an analysis. Consequently, the arguments of [2] and subsequent research are cast in a new light. Focusing on arbitrary, but finite N not only simplifies the analysis, but allows us to freely combine well-known results from convex, linear and integer linear optimization theory to obtain results applicable to practical systems. Last but not least, the presented approach shows that while the results of [3], [4] on the duality gap of non-convex optimization problems are valuable in general, they are not necessary for the case of multiuser multicarrier RAPs.

The primary contribution of this work is to increase the understanding of dual methods and their performance with re-

Manuscript received August 12, 2013; revised November 20, 2013 and January 27, 2014; accepted February 2, 2014. The associate editor coordinating the review of this paper and approving it for publication was Z. Han.

This work was supported by the UMIC research cluster at RWTH Aachen University.

The authors are with the Institute for Theoretical Information Technology, RWTH Aachen University, Aachen, Germany (e-mail: {goertzen, schmeink}@umic.rwth-aachen.de).

Digital Object Identifier 10.1109/TWC.051414.131480

spect to resource allocation problems. The problems appearing in this paper are meant to serve as blueprints for the problems encountered in practical systems. We formally prove that the dual solution has a near-discrete structure and calculate absolute and relative performance bounds, which simplifies the implementation of dual methods with respect to numerical precision and algorithm termination. We furthermore show that the performance of a dual method strongly depends on the rounding method used to obtain a discrete solution, which is an often overlooked step of high practical importance.

The paper is organized as follows. Section II covers the system model and notation. Also included is a detailed description of the section structure of Section III and Section IV. In Section III, the sum rate maximization problem with a global power budget and the sum power minimization problem with a global rate demand are analyzed. Section IV covers the sum rate maximization problem with individual power budgets and the sum power minimization problem with individual rate demands. Each of the aforementioned sections includes a performance evaluation of dual methods for the respective problems. Section V concludes the paper.

II. Preliminaries

We consider a wireless communication system with K users and N orthogonal subcarriers, in which each subcarrier can only be used by at most one user. Let $p_{k,n}$ denote the transmit power spent for user k on subcarrier n. Then, user k's data rate on subcarrier n is given as $r_{k,n}(p_{k,n})$ for a rate utility function $r_{k,n}\colon \mathbb{R}_+ \to \mathbb{R}_+$ incorporating channel gain information and other factors. The classical example of a concave rate function given channel gain $g_{k,n}$ is $r_{k,n}(p_{k,n}) = \log_2(1 + g_{k,n}p_{k,n})$, which is based on the Shannon bound [14]. With the appropriate scaling factors, this function becomes a good approximation of a multitude of practical modulation and coding schemes (MCSs) which correspond to discontinuous rate functions [15]. For the purpose of this work, we assume all rate functions to be continuous but not necessarily differentiable. Furthermore, we assume the following properties to hold for all k and n:

1) $r_{k,n}$ is concave as a function of $p_{k,n}$,
2) $r_{k,n}(0) = 0$,
3) $\lim_{t\to\infty} r_{k,n}(t)/t = 0$.

The Shannon rate formula above satisfies all of these properties, as do most variations based on scaling factors. Note that the term $\lim_{t\to\infty} r_{k,n}(t)/t$ in 3) denotes the slope of the line passing through $r_{k,n}(0) = 0$ and $r_{k,n}(t)$ as t approaches infinity. For linear functions, it coincides with their slope, and it is 0 for bounded functions. Any rate function is easily modified to satisfy 3) as the achievable rates of practical systems are always bounded.

In this paper, we analyze four resource allocation problems which are formally introduced in the following sections. They are the sum rate maximization problem (SRMP), the sum rate maximization problem with individual budgets (SRMPI), the sum power minimization problem (SPMP), and the sum power minimization problem with individual demands (SPMPI). Refer to Table I for an overview. For problems with individual requirements, K constraints govern the power budget or rate demand of each user. For problems with a global requirement, a single constraint ensures that the sum power or sum rate of all users satisfies a given bound. Section III covers resource allocation under a global constraint. This includes an analysis of the SRMP in full detail, followed by an analysis of the SPMP. The section concludes with a joint performance evaluation for both problems. Section IV covers the problem of resource allocation under user-individual constraints. The analysis of the SRMPI and the SPMPI is presented in this section, followed again by a performance evaluation.

TABLE I
An Overview of the Four Problems Analyzed in This Paper.

Section	Problem	Objective	Constraint Set Type	Size
III-A	SRMP	sum rate max.	global power budget	1
III-B	SPMP	sum power min.	global rate demand	1
IV-A	SRMPI	sum rate max.	individual power budgets	K
IV-B	SPMPI	sum power min.	individual rate demands	K

For ease of reference and comparison, the analysis of each RAP in the aforementioned sections is structured in the same way. Note that analogous proofs in the analysis of later sections are omitted to avoid repetition. We provide a short overview of this structure in the following list:

1) **Formulation –** a multiple-choice knapsack-based problem formulation of the RAP with continuous power variables $\{p_{k,n}\}_{k,n}$ and binary allocation variables $\{x_{k,n}\}_{k,n}$ is introduced.
2) **Relaxation –** the binary condition of the allocation variables is relaxed to obtain an RAP with subcarrier sharing.
3) **Convexity –** the convexity of the relaxed problem is shown.
4) **Duality –** the concurrent dual problem of the above problems is introduced.
5) **Performance –** a bound on the number of nonbinary components in solutions to the relaxed problem is presented, and it is shown how to obtain feasible roundings[1] and performance guarantees based on this bound.

The novelty of the presented approach lies in the formulation used to describe the RAPs. It is inspired by so-called multiple-choice knapsack problems in integer linear programming [16]. Different from formulations that restrict users' power levels to nonconvex domains, the unique subcarrier usage constraint is made explicit by introducing additional binary variables that govern which subcarriers are assigned to which user. Doing so results in a nonconvex problem formulation with nonconvex constraint set of size 1 in the case of a global constraint, and size K in the case of individual constraints. Throughout this paper, modified and transformed problems related to each of the four original RAPs appear. Denoting the original RAP by (P), which we assume to be feasible, they are introduced in the following summary:

As an RAP with a fixed and finite number of subcarriers (as opposed to $N \to \infty$), problem (P) does not satisfy the time-sharing property of [2]. Linearly relaxing the integer subcarrier allocation constraint yields (P-S), which corresponds to a system with subcarrier sharing. By definition, (P-S) satisfies

[1]This formal approach fails in the case of the SPMPI. Refer to Section IV-C for details.

the time-sharing property. Thus, it is known from [2] that there is zero duality gap between (P-S) and its dual problem, denoted (P-D). In this paper, we show how to obtain equivalent convex formulations, denoted (P-C), for each of the linear relaxations (P-S) of the examined RAPs. This extends the work of [2] as zero duality gap then follows from checking for a constraint qualification like Slater's condition.

With the performance analysis described in 5), this paper further contributes to the question of whether dual methods for RAPs can be considered optimal. We present a bound for the number of nonbinary components that occur in solutions to (P-S), which can be transformed into solutions to the dual problem (P-D), and vice versa. By showing that feasible points of (P) can be obtained from these solutions, it is possible to give a performance guarantee for dual methods. To the best of the authors' knowledge, this is the first performance analysis which does not depend on the asymptotic arguments that provide the basis for the optimality claims of [2].

Notation

We introduce shorthand notation for a few recurring formulas. For a rate function $r_{k,n}$, we define the conjugate functions $r^*_{k,n}(\lambda) = \sup_{p_{k,n}} \{r_{k,n}(p_{k,n}) - \lambda p_{k,n}\}$ and $r^\circ_{k,n}(\lambda) = \inf_{p_{k,n}} \{p_{k,n} - \lambda r_{k,n}(p_{k,n})\}$. See Fig. 1 for a visualization. Given slope λ, $r^*_{k,n}(\lambda)$ is the smallest value such that the line $\lambda p_{k,n} + r^*_{k,n}(\lambda)$ is greater or equal to $r_{k,n}(p_{k,n})$ for all $p_{k,n}$. For $\lambda > 0$, it holds that $r^\circ_{k,n}(\lambda) = -\lambda r^*(1/\lambda)$.

For a set of real-valued variables $\{x_{k,n}\}_{k,n}$ and a set $\mathcal{S}$, we define $c_\mathcal{S} = |\{(k,n)\,|\,x_{k,n} \in \mathcal{S}\}|$. The term $c_\mathcal{S}$ counts the number of variables with values within the set $\mathcal{S}$.

III. Resource Allocation under a Global Constraint

A. Sum Rate Maximization with a Global Power Budget

Given a global power constraint $p^{\max} > 0$ and rate functions $r_{k,n}$, the SRMP is an optimization problem with variables $\{p_{k,n}\}_{k,n}$ and $\{x_{k,n}\}_{k,n}$. The binary variable $x_{k,n}$ has a value of one if user k is using subcarrier n, and a value of zero otherwise.

$$\text{(SRMP)} \quad \underset{p_{k,n}\ge 0,\, x_{k,n}\in\{0,1\}}{\text{maximize}} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} r_{k,n}(p_{k,n})$$

$$\text{subject to} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} p_{k,n} \le p^{\max}, \qquad \text{(GPB)}$$

$$\sum_{k=1}^{K} x_{k,n} = 1, \quad n = 1,\ldots,N. \qquad \text{(MC)}$$

The above formulation detaches the problem of allocating subcarriers and power by modeling each with separate variables. The global power budget constraint is given by (GPB), and the multiple-choice constraint (MC) governs unique subcarrier usage.

One has to take special care when applying duality theory to nonconvex problems. Among other things, it is not guaranteed that the dual optimum can be used to obtain a feasible primal solution. Furthermore, the duality gap, the difference between primal and dual objectives, is not guaranteed to vanish. For convex problems, both of these properties hold under mild conditions.

While (MC) is affine, (SRMP) is not a convex problem. As is well known, the product of two variables is neither convex nor concave in general. Thus, the objective of (SRMP) is not concave and (GPB) is not convex. Furthermore the binary condition $x_{k,n} \in \{0,1\}$ is not even continuous. Hence, a necessary requirement for convexity is a relaxation of the binary condition. This constitutes the first step of our analysis.

The convex relaxation of the binary condition $x_{k,n} \in \{0,1\}$ is given by $0 \le x_{k,n} \le 1$. This relaxation has a sensible and meaningful interpretation. In the relaxed problem formulation, instead of unique subcarrier usage, users can share resource blocks as if modeling a time-sharing system. The nonbinary factors $x_{k,n}$ appropriately scale both rate output and power demand. We refer to this relaxed problem as the SRMP with subcarrier sharing. It can be formulated as follows:

$$\text{(SRMP-S)} \quad \underset{p_{k,n}\ge 0,\, x_{k,n}\ge 0}{\text{maximize}} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} r_{k,n}(p_{k,n})$$
$$\text{subject to} \quad \text{(GPB) and (MC).}$$

Note that the constraint $x_{k,n} \le 1$ is redundant given (MC). The above formulation is still nonconvex due to the nonconcave objective and nonconvex (GPB). However, there exists an equivalent convex formulation.

Proposition 1. *There exists an equivalent convex formulation of (SRMP-S). It is given by*

$$\textit{(SRMP-C)} \quad \underset{q_{k,n}\ge 0,\, x_{k,n}\ge 0}{\textit{maximize}} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} r_{k,n}(q_{k,n}/x_{k,n})$$
$$\textit{subject to} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} q_{k,n} \le p^{\max} \textit{ and (MC),}$$

in which we define $x_{k,n} r_{k,n}(q_{k,n}/x_{k,n}) = 0$ for $x_{k,n} = 0$.

Proof: We prove the convexity of (SRMP-C) by showing that the objective is a sum of concave functions. The well-known perspective function of r for $x > 0$ is defined as $f(q,x) = x \cdot r(q/x)$ [17]. The perspective function of a concave function is concave. This property persists when continuously extending f towards the closure of its domain [18]. The continuous extension of f to include $x = 0$ is given by

$$f(q,0) = \lim_{x\to 0} x \cdot r\left(\frac{q}{x}\right) \overset{t=\frac{q}{x}}{=} \lim_{t\to\infty} \frac{q}{t} \cdot r(t) = q \cdot \lim_{t\to\infty} \frac{r(t)}{t},$$

which equals 0 for the concave rate functions in this paper according to property 3) in Section II. Hence, the functions in the objective are continuously extended (or closed) perspective functions of $r_{k,n}$ and thus concave. This proves the convexity of (SRMP-C).

(SRMP-C) is obtained from (SRMP-S) by the substitution of variable $q_{k,n} = x_{k,n} p_{k,n}$. This substitution is not one-to-one, but each solution to (SRMP-S) is transformed to a solution to (SRMP-C) and vice versa, which can be shown as follows: Let $\{x_{k,n}, p_{k,n}\}_{k,n}$ denote a solution to (SRMP-S). Because all occurences of $p_{k,n}$ are multiplied by $x_{k,n}$ in the objective and constraints of (SRMP-S), we can set $p_{k,n} = 0$ whenever $x_{k,n} = 0$ without loss of optimality. As a result,

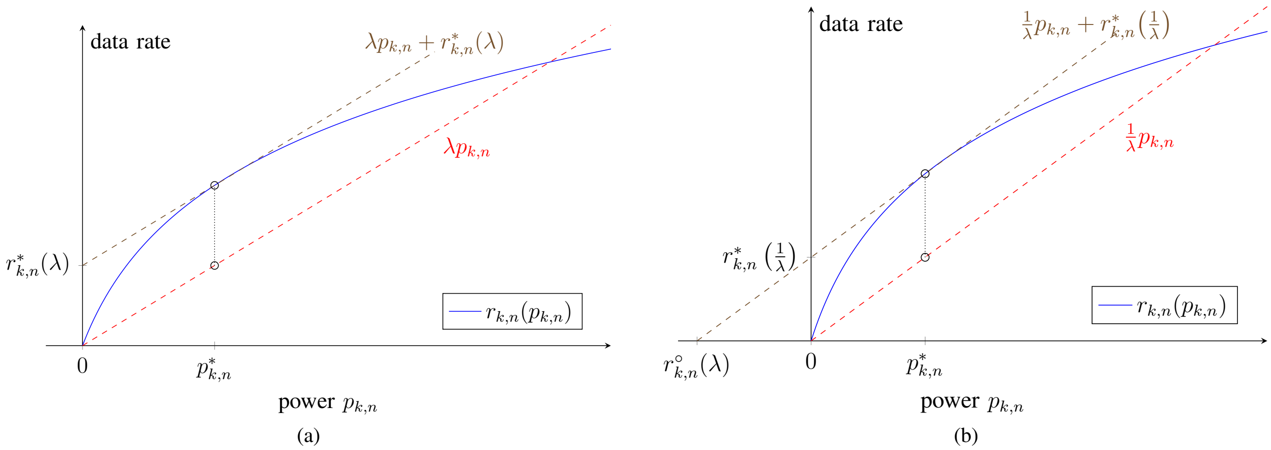


Fig. 1. Visualization of $r^*_{k,n}(\lambda)$ and $r^\circ_{k,n}(\lambda)$, respectively. (a) The value of $r^*_{k,n}(\lambda)$ is the y-intercept of the tangent with slope λ when approaching the function $r_{k,n}$ from above. (b) The value of $r^\circ_{k,n}(\lambda)$ is the x-intercept of the tangent with slope $1/\lambda$ when approaching $r_{k,n}$ from above.

the mapping $\{x_{k,n}, p_{k,n}\}_{k,n} \to \{x_{k,n}, q_{k,n}\}_{k,n}$ is one-to-one. The inverse mapping is given by the substitution

$$p_{k,n} = \begin{cases} q_{k,n}/x_{k,n}, & \text{if } x_{k,n} > 0, \\ 0, & \text{else.} \end{cases}$$

The above substitutions map solutions of (SRMP-S) to solutions of (SRMP-C) and vice versa. This shows that (SRMP-S) and (SRMP-C) are equivalent. ■

Note that the duality gap between RAPs which satisfy the so-called time-sharing property and their dual problems has been shown to be zero [2]. Relaxing $x_{k,n} \in \{0, 1\}$ allows for arbitrary subcarrier sharing and thus yields a problem that trivially satisfies the time-sharing property. Hence, the results of [2] imply that (SRMP-S) and its dual have zero duality gap. By comparison, the above proposition shows that there exists an equivalent convex formulation (SRMP-C). This is a stronger result as a duality gap of zero then follows under the mild condition that one of many constraint qualifications hold. The convexity of (SRMP-C) and thus (SRMP-S) is going to prove advantageous in the following analysis and eliminates the need for the asymptotic arguments of [2].

While the dual problems of equivalent problems do not have to coincide in general, the dual problems of (SRMP-S) and (SRMP-C) are identical. Furthermore, strong duality holds, i.e., there is zero duality gap between the equivalent primal problems and their dual problem (SRMP-D). We show the above and state (SRMP-D) in the following proposition.

Proposition 2. *The dual problems of (SRMP), (SRMP-S) and (SRMP-C) coincide. The dual problem is*

$$\text{(SRMP-D)} \quad \underset{\lambda}{\text{minimize}} \quad \lambda p^{\max} + \sum_{n=1}^{N} \max_k r^*_{k,n}(\lambda)$$
$$\text{subject to} \quad \lambda \geq 0.$$

Problems (SRMP-S) and (SRMP-D) have zero duality gap and a solution to (SRMP-S) can be obtained from the dual optimum.

Proof: We prove the last statement first. As shown in Prop. 1, (SRMP-S) and the convex problem (SRMP-C) are equivalent. In order to apply Slater's theorem, Slater's condition has to hold [17]. Let $\{q_{k,n}, x_{k,n}\}_{k,n}$ denote a feasible point of (SRMP-C), the feasibility of which follows from the feasibility of (SRMP). If $\sum_{k=1}^{K}\sum_{n=1}^{N} q_{k,n} < p^{\max}$, this point is an interior point of its domain and Slater's condition is satisfied. Otherwise, the condition can be satisfied by decreasing one of the $q_{k,n} > 0$, which does not affect feasibility. Slater's theorem now implies that strong duality holds and a solution to (SRMP-C), and, equivalently, (SRMP-S), can be obtained from the dual optimum.

As (SRMP) and (SRMP-S) only differ in their domain ($x_{k,n} \in \mathbb{N}_0$ and $x_{k,n} \in \mathbb{R}_+$, respectively), their dual problems coincide. It remains to be shown that the dual problems of (SRMP-S) and (SRMP-C) coincide. For (SRMP-S), the dual function is

$$\begin{aligned}
&g_{\text{(SRMP-S)}}(\lambda, \mu_1, \ldots, \mu_N) = \sup_{\{x_{k,n}, p_{k,n}\}} \Bigg\{ \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} r_{k,n}(p_{k,n}) \\
&+ \lambda \left(p^{\max} - \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} p_{k,n} \right) + \sum_{n=1}^{N} \mu_n \left(1 - \sum_{k=1}^{K} x_{k,n} \right) \Bigg\} \\
&= \lambda p^{\max} + \sum_{n=1}^{N} \mu_n + \sum_{k=1}^{K}\sum_{n=1}^{N} \sup_{x_{k,n}} \Bigg\{ x_{k,n} \\
&\qquad \cdot \left(\sup_{p_{k,n}} \{ r_{k,n}(p_{k,n}) - \lambda p_{k,n} \} - \mu_n \right) \Bigg\} \\
&= \lambda p^{\max} + \sum_{n=1}^{N} \mu_n + \sum_{k=1}^{K}\sum_{n=1}^{N} \sup_{x_{k,n}} \{ x_{k,n} (r^*_{k,n}(\lambda) - \mu_n) \} \qquad (1) \\
&= \begin{cases} \lambda p^{\max} + \sum_{n=1}^{N} \mu_n, & \text{if } \mu_n \geq r^*_{k,n}(\lambda) \text{ for all } k, n, \\ \infty, & \text{else.} \end{cases} \qquad (2)
\end{aligned}$$

Conversely, the dual function of (SRMP-C) is

$$g_{\text{(SRMP-C)}}(\lambda, \mu_1, \ldots, \mu_N) = \lambda p^{\max} + \sum_{n=1}^{N} \mu_n + \sum_{k=1}^{K}\sum_{n=1}^{N} \sup_{x_{k,n}} \left\{ \sup_{q_{k,n}} \{ x_{k,n}(r_{k,n}(q_{k,n}/x_{k,n}) - \mu_n) - \lambda q_{k,n} \} \right\}. \qquad (3)$$

Note that for $x_{k,n} = 0$, the inner supremum in (3) equals $\sup_{q_{k,n}} \{-\lambda q_{k,n}\} = 0$. For $x_{k,n} > 0$, we can apply *the substitution $p_{k,n} = q_{k,n}/x_{k,n}$ and obtain*

$$\begin{aligned}
&\sup_{q_{k,n}} \{x_{k,n}(r_{k,n}(q_{k,n}/x_{k,n}) - \mu_n) - \lambda q_{k,n}\} \\
&= \sup_{p_{k,n}} \{x_{k,n}(r_{k,n}(p_{k,n}) - \mu_n) - \lambda x_{k,n} p_{k,n}\} \\
&= x_{k,n} \cdot \sup_{p_{k,n}} \{r_{k,n}(p_{k,n}) - \lambda p_{k,n} - \mu_n\} \\
&= x_{k,n} \left(r^*_{k,n}(\lambda) - \mu_n\right). \qquad (4)
\end{aligned}$$

As the above evaluates to 0 for $x_{k,n} = 0$, this formulation holds in both cases. Plugging (4) into (3) yields (1) and we conclude $g_{\text{(SRMP-S)}} = g_{\text{(SRMP-C)}}$. As shown in (2), $\mu_n \geq r^*_{k,n}(\lambda)$ has to hold for all k and n to ensure $g_{\text{(SRMP-S)}} < \infty$. The minimal and therefore optimal choice is to set $\mu_n = \max_k r^*_{k,n}(\lambda)$ and (SRMP-D) is obtained. ■

Prop. 1 and Prop. 2 show that a solution to (SRMP-S) can be obtained from the dual optimum. However, this does not mean that the output is a feasible point of the original problem. If this were the case it would not only be feasible, but optimal for (SRMP), which might be true in special cases, but does not hold in general.

The performance of a dual method depends on the duality gap between (SRMP) and (SRMP-C). Prop. 2 shows that the duality gap is equal to the difference between the optimal values of the discrete problem (SRMP) and the continuous problem (SRMP-S). An upper bound for this difference is an upper bound for the duality gap and thus a bound on the overall performance of a dual method.

We present and analyze the SPMP in the next subsection before jointly evaluating the performance of dual methods for both problems in Section III-C.

B. Sum Power Minimization with a Global Rate Demand

Given a global rate demand $r^{\min} > 0$, the SPMP can be formulated as follows:

$$\begin{aligned}
\text{(SPMP)} \quad & \underset{p_{k,n}\geq 0,\, x_{k,n}\in\{0,1\}}{\text{minimize}} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} p_{k,n} \\
\text{subject to} \quad & \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} r_{k,n}(p_{k,n}) \geq r^{\min}, \qquad \text{(GRD)} \\
& \sum_{k=1}^{K} x_{k,n} = 1, \quad n = 1, \ldots, N. \qquad \text{(MC)}
\end{aligned}$$

Here, (GRD) denotes the global rate demand constraint. As before, it is necessary to relax $x_{k,n} \in \{0, 1\}$ to obtain a convex problem. The SPMP with subcarrier sharing is given by

$$\begin{aligned}
\text{(SPMP-S)} \quad & \underset{p_{k,n}\geq 0,\, x_{k,n}\geq 0}{\text{minimize}} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} p_{k,n} \\
& \text{subject to} \quad \text{(GRD) and (MC).}
\end{aligned}$$

We obtain the following result:

Proposition 3. *The dual problems of (SPMP) and (SPMP-S) coincide. With $r^\circ_{k,n}(\lambda) = \inf_{p_{k,n}} \{p_{k,n} - \lambda r_{k,n}(p_{k,n})\}$, the dual problem is*

$$\begin{aligned}
\textit{(SPMP-D)} \quad & \underset{\lambda}{\textit{maximize}} \quad \lambda r^{\min} + \sum_{n=1}^{N} \min_k r^\circ_{k,n}(\lambda) \\
& \textit{subject to} \quad \lambda \geq 0.
\end{aligned}$$

Problems (SPMP-S) and (SPMP-D) have zero duality gap and a solution to (SPMP-S) can be obtained from the dual optimum.

Proof: Applying the transformation $q_{k,n} = x_{k,n}p_{k,n}$ to (SPMP-S) yields an equivalent convex problem, which is shown analogously to Prop. 1. The dual function for (SPMP-S) is

$$\begin{aligned}
g_{\text{(SPMP-S)}}(\lambda, \mu_1, \ldots, \mu_N) &= \inf_{\{x_{k,n},\, p_{k,n}\}} \Bigg\{ \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} p_{k,n} \\
&\quad + \lambda \left(r^{\min} - \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} r_{k,n}(p_{k,n}) \right) \\
&\quad + \sum_{n=1}^{N} \mu_n \left(1 - \sum_{k=1}^{K} x_{k,n} \right) \Bigg\} \\
&= \lambda r^{\min} + \sum_{n=1}^{N} \mu_n + \sum_{k=1}^{K}\sum_{n=1}^{N} \inf_{x_{k,n}} \Bigg\{ x_{k,n} \\
&\quad \cdot \left(\inf_{p_{k,n}} \{p_{k,n} - \lambda r_{k,n}(p_{k,n})\} - \mu_n \right) \Bigg\} \\
&= \lambda r^{\min} + \sum_{n=1}^{N} \mu_n + \sum_{k=1}^{K}\sum_{n=1}^{N} \inf_{x_{k,n}} \left\{ x_{k,n} \left(r^\circ_{k,n}(\lambda) - \mu_n \right) \right\} \\
&= \begin{cases} \lambda r^{\min} + \sum_{n=1}^{N} \mu_n, & \textit{if } \mu_n \leq r^\circ_{k,n}(\lambda) \textit{ for all } k, n, \\ -\infty, & \textit{else.} \end{cases}
\end{aligned}$$

To ensure $g_{\text{(SRMPI-S)}} < \infty$, $\mu_n \leq r^\circ_{k,n}(\lambda)$ has to hold for all k and n. The maximal and thus optimal choice is to set $\mu_n = \min_k r^\circ_{k,n}(\lambda_k)$ and (SPMP-D) is obtained. We omit the remaining proof as it is otherwise analogous to the proof of Prop. 2. ■

C. Performance under a Global Constraint

As previously mentioned, the convexity of (SRMP-S) and (SPMP-S) means that the performance gap of a dual method is given by the difference in objective between the discrete original problem and its continuous relaxation. In order to bound this gap, we show that relaxed solutions can be transformed, or rounded, towards a feasible point of the original problem. The key to obtaining a bound on the absolute performance is the fact that the number of nonbinary components in solutions to (SRMP-S) and (SPMP-S) can be bounded.

We cite a result for linear optimization problems, or linear programs (LPs) that is based on the geometry of polyhedra. As such, it is related to the proof of the Shapley-Folkman Theorem [19, App. I] and plays a crucial role for LPs in general and the simplex algorithm [20, Ch. I.2.3] in particular. After some simplifications and minor notational changes we obtain from [20]:

Corollary 1. *If an LP of the form* $\max_{\boldsymbol{x}} \left\{ \boldsymbol{r}^T\boldsymbol{x} \,\middle|\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\, \boldsymbol{x} \geq \boldsymbol{0} \right\}$ *is feasible, it has a solution with at most* $\operatorname{rank}(\boldsymbol{A})$ *nonzero components of* $\boldsymbol{x}$.

Proof: This follows from Def. 3.1 and Thm. 3.5 in [20, Ch. I.2.3]. ∎

The above corollary can be applied to bound the number of nonbinary components in solutions to (SRMP-S) and (SPMP-S), respectively:

Proposition 4. *Let (P-S) denote a feasible instance of either (SRMP-S) or (SPMP-S). Then, there exists a solution* $\{x_{k,n}, p_{k,n}\}_{k,n}$ *to (P-S) that satisfies*

1) $c_{\{1\}} := |\{(k,n) \,|\, x_{k,n} = 1\}| \geq N - 1$,
2) $c_{(0,1)} := |\{(k,n) \,|\, x_{k,n} \in (0,1)\}| \leq 2$.

In other words, there is at most one subcarrier for which $\{x_{k,n}\}_{k,n}$ *has nonbinary components. Furthermore, the number of nonbinary components is at most two.*

Proof: Let $\{\tilde{x}_{k,n}, p_{k,n}\}_{k,n}$ *denote a solution to (P-S). Fixing the power values to* $\{p_{k,n}\}_{k,n}$*, we obtain a feasible LP with optimization variables* $\{x_{k,n}\}_{k,n}$*. We denote this LP with fixed power values by (P-F). Note that each solution* $\{x_{k,n}\}_{k,n}$ *to (P-F) maps to a solution* $\{x_{k,n}, p_{k,n}\}_{k,n}$ *to (P-S). To apply Cor. 1, the one inequality contraint of (P-F) has to be transformed to an equality constraint. This is achieved by adding a nonnegative slack variable to the left side of (GPB) or the right side of (GRD), respectively. As a result, (P-F) is an LP of the form of Cor. 1 with a constraint matrix* $\boldsymbol{A}$ *with* $N+1$ *rows, corresponding to the* N *multiple-choice constraints (MC) and (GPB) or (GRD), respectively.*

Based on Cor. 1, there exists a solution $\{x_{k,n}\}_{k,n}$ *to (P-F) satisfying* $c_{(0,1]} := |\{(k,n) \,|\, x_{k,n} \in (0,1]\}| \leq \operatorname{rank}(\boldsymbol{A}) \leq N+1$*, as the rank of* $\boldsymbol{A}$ *is upper bounded by its row dimension. Because each subcarrier with nonbinary components has to include at least two nonbinary* $x_{k,n}$ *to satisfy (MC), it holds that* $2\left(N - c_{\{1\}}\right) \leq c_{(0,1)}$*. Together, it follows that*

$$2\left(N - c_{\{1\}}\right) \leq c_{(0,1)} = c_{(0,1]} - c_{\{1\}} \leq N + 1 - c_{\{1\}}. \quad (5)$$

Subtracting $N - c_{\{1\}}$ *from both sides yields* $N - c_{\{1\}} \leq 1$*, which is equivalent to 1) and yields 2) from* (5). ∎

The above results can be summarized as follows: Introducing subcarrier sharing to one of the RAPs with a global constraint results in a relaxed solution in which at most one subcarrier is shared between two users. The remaining $N-1$ subcarriers are used exclusively by one user. Somewhat counterintuitively, the optimal operating procedure in a system with full subcarrier sharing is local subcarrier sharing: at most one subcarrier is shared.

Prop. 4 allows for a quick informal estimate of the expected relative performance. Allocating a shared subcarrier to one of the users is going to lead to a minor decrease in rate on a single subcarrier. This performance loss can be expected to be of a small degree $\rho \ll 1$ compared to the overall data rate provided by this subcarrier. Looking at the relative performance, there are N subcarriers which provide data rate. Assuming that there is no unbalance or bias within the subcarriers, this means that the relative performance loss is given by $\rho/N < 1/N$ which is very close to 0 for even a moderate number of subcarriers. The remainder of this section is devoted to formalizing the above estimate.

To obtain a formal performance estimate, it is necessary to derive a feasible point for the original problem from the solution to the RAP with subcarrier sharing. This is achieved by assigning the one subcarrier being shared to exactly one user. In the following, we assume w.l.o.g. that the subcarrier n is shared between users $k=1$ and $k=2$.

Let $\{p_{k,n}, x_{k,n}\}_{k,n}$ be a solution to (SRMP-S) satisfying the bounds of Prop. 4. In the case that $\{x_{k,n}\}_{k,n}$ is binary, this solution is also optimal for (SRMP) because (SRMP-S) is a relaxation of (SRMP). Otherwise, it has exactly two nonbinary components which can be rounded to obtain two feasible points of (SRMP). The roundings are given by the mappings

$$\begin{aligned}(x_{1,n}, p_{1,n}, x_{2,n}, p_{2,n}) &\mapsto (1, x_{1,n}p_{1,n} + x_{2,n}p_{2,n}, 0, 0) \quad \text{and} \\ (x_{1,n}, p_{1,n}, x_{2,n}, p_{2,n}) &\mapsto (0, 0, 1, x_{1,n}p_{1,n} + x_{2,n}p_{2,n}),\end{aligned}$$

which is verified by checking that the right hand side satisfies (GPB), (MC) and $x_{k,n} \in \{0,1\}$ for all k and n.

Let $\{p_{k,n}, x_{k,n}\}_{k,n}$ be a solution to (SPMP-S) satisfying the bounds of Prop. 4. In the case that $\{x_{k,n}\}_{k,n}$ is binary, this solution is optimal for (SPMP) as (SPMP-S) relaxes (SPMP). Otherwise, it has exactly two nonbinary components which can be rounded to obtain a feasible point of (SPMP).

Denote by $\tilde{r} = x_{1,n}r_{1,n}(p_{1,n}) + x_{2,n}r_{2,n}(p_{2,n})$ the rate obtained on subcarrier n by the optimal shared solution. The power values for the two possible mappings are given by

$$\begin{aligned}\psi_1(\tilde{r}) &:= \inf\left\{p \,|\, r_{1,n}(p) \geq \tilde{r}\right\} \quad \text{and} \\ \psi_2(\tilde{r}) &:= \inf\left\{p \,|\, r_{2,n}(p) \geq \tilde{r}\right\}.\end{aligned}$$

As $\max\{r_{1,n}(p_{1,n}), r_{2,n}(p_{2,n})\} \geq \tilde{r}$ holds, at least one of these terms is finite. Then, the mappings

$$\begin{aligned}(x_{1,n}, p_{1,n}, x_{2,n}, p_{2,n}) &\mapsto (1, \psi_1(\tilde{r}), 0, 0) \quad \text{and} \\ (x_{1,n}, p_{1,n}, x_{2,n}, p_{2,n}) &\mapsto (0, 0, 1, \psi_2(\tilde{r}))\end{aligned}$$

yield feasible points for $\psi_1(\tilde{r}) < \infty$ and $\psi_2(\tilde{r}) < \infty$, respectively.

It has been shown that applying a dual method to (SRMP) or (SPMP) results in a subcarrier assignment that is unique except for at most one shared subcarrier. As the goal is to analyze the performance of dual methods, we ignore the potential for additional optimization steps to improve on this approach. Note that the power distribution given by the above mappings are feasible, but not necessarily optimal for the corresponding assignment. However, once a subcarrier assignment is fixed, the optimal power distribution is given by the well-known waterfilling procedure.

In order to give a performance guarantee, the rate loss/power increase of the above roundings has to be measured. We drop the index n of the shared subcarrier for readability and assume the users to be ordered such that $r_1(p_1) > r_2(p_1)$ and $r_2(p_2) > r_1(p_2)$ holds. Define $\tilde{p} = x_1p_1 + x_2p_2$. Then, the rate loss L_{SRMP} suffered from allocating the shared subcarrier to the user with the higher rate is given by

$$L_{\text{SRMP}} = x_1r_1(p_1) + x_2r_2(p_2) - \max\left\{r_1(\tilde{p}),\, r_2(\tilde{p})\right\}.$$

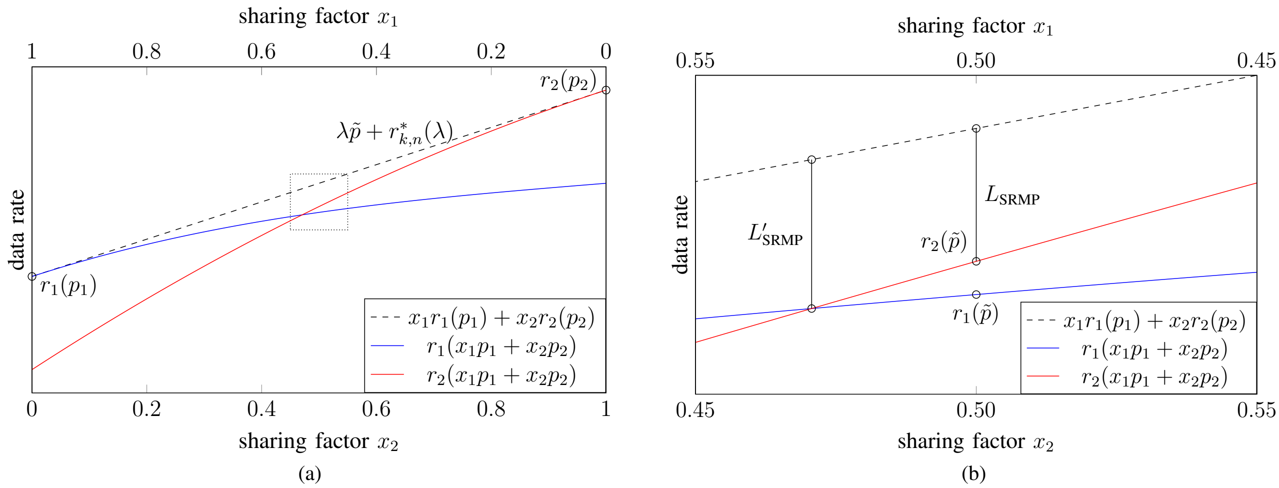


Fig. 2. Visualization of the absolute performance loss induced by rounding the SRMP. (a) Depicted are the rate functions of both users for $\tilde{p} = x_1 p_1 + x_2 p_2$ in the interval $[p_1, p_2]$ in which performance is lost due to rounding. (b) Zooming in on $\tilde{p}$ for the exemplary $x_1 = x_2 = \frac{1}{2}$ shows that L_{SRMP} depends on x_1 and x_2, whereas L'_{SRMP} is an independent upper bound.

The term L_{SRMP} depends on the value of x_1 and $x_2 = 1 - x_1$. A bound that does not depend on x_1 is given by the following worst-case analysis:

$$\begin{aligned} &L'_{\text{SRMP}} \\ &= \max_{0 \le x_1 \le 1} \left\{ x_1 r_1(p_1) + x_2 r_2(p_2) - \max\left\{ r_1(\tilde{p}),\, r_2(\tilde{p}) \right\} \right\} \quad (6) \\ &= \max_{p_1 \le \tilde{p} \le p_2} \left\{ \lambda\tilde{p} + r^*_{k,n}(\lambda) - \max\left\{ r_1(\tilde{p}), r_2(\tilde{p}) \right\} \right\} \quad (7) \end{aligned}$$

In (6), $x_2 = 1 - x_1$ and $\tilde{p} = x_1 p_1 + x_2 p_2$ depend on x_1. In (7), the term $\lambda\tilde{p} + r^*_{k,n}(\lambda)$ describes the linear function which is tangent to r_1 at p_1 and to r_2 at p_2, respectively. Note that L'_{SRMP} describes the maximum gap between the function $\max\{r_1, r_2\}$ and the convex closure of its graph. This reinforces that the duality gap is a measure for non-convexity (or non-concavity in this case). See Fig. 2 for a visualization of the performance loss bounds L_{SRMP} and L'_{SRMP}.

As the objective of the SPMP is power minimization, the performance loss of a rounding has to be measured by the additional power spent to achieve the same sum rate as the shared solution. As the shared solution shares both rate and power between users, the power spent on the shared solution equals $\tilde{p} = x_1 p_1 + x_2 p_2$. This value has to be compared to the power spent on one of the feasible solutions to (SPMP). An optimal rounding consists of choosing the subcarrier assignment which requires the least power, given by the minimum of $\psi_1(\tilde{r})$ and $\psi_2(\tilde{r})$. The additional power spent computes to

$$L_{\text{SPMP}} = \min\{\psi_1(\tilde{r}), \psi_2(\tilde{r})\} - x_1 p_1 - x_2 p_2.$$

Maximizing this term with respect to these sharing factors constitutes a worst-case analysis:

$$\begin{aligned} L'_{\text{SPMP}} &= \max_{0 \le x_1 \le 1} \left\{ \min\{\psi_1(\tilde{r}), \psi_2(\tilde{r})\} - x_1 p_1 - x_2 p_2 \right\} \quad (8) \\ &= \max_{r_1(p_1) \le \tilde{r} \le r_2(p_2)} \left\{ \min\{\psi_1(\tilde{r}), \psi_2(\tilde{r})\} - \lambda\tilde{r} - r^\circ_{k,n}(\lambda) \right\}. \quad (9) \end{aligned}$$

In (8), the term $\tilde{r} = x_1 r_1(p_1) + x_2 r_2(p_2)$ is a function of x_1. The equivalent formulation (9) shows that the performance loss is given by the maximum difference between the function $\min\{\psi_1(\tilde{r}), \psi_2(\tilde{r})\}$ and the coinciding tangent function given by $p = \lambda\tilde{r} + r^\circ_{k,n}(\lambda)$.

See Fig. 3 for a visualization of the performance loss bounds L_{SPMP} and L'_{SPMP}. The bound L'_{SPMP} describes the maximum gap between the function $\min\{\psi_1, \psi_2\}$ and the convex closure of its graph. This result resembles that of L'_{SRMP}. Note that $\min\{\psi_1, \psi_2\}$ is the inverse function to $\max\{r_1, r_2\}$. As function inversion corresponds to graph transposition, the graph of $\min\{\psi_1, \psi_2\}$ is the transposed graph of $\max\{r_1, r_2\}$. Hence, the performance bound L'_{SPMP} corresponds to the *horizontal* gap between $\max\{r_1, r_2\}$ and its convex closure. See Fig. 4 for a direct comparison of the performance bounds for the SRMP and the SPMP.

We have presented two performance bounds L'_{SRMP} and L'_{SPMP}. However, both of these bounds rely on the information which subcarrier is shared between which users. In order to obtain an a-priori performance estimate, we compute the global maximum value of the above bounds. Define

$$r_n(p) = \max_k r_{k,n}(p) \quad \text{and} \quad \psi_n(r) = \inf\{p \,|\, r_n(p) \ge r\}.$$

Then, the data rate of the convex closure of r_n at power level p is given by the best possible tangential approximation $r_n^{\text{conv}}(p) = \inf_\lambda\{\lambda p + r^*_n(\lambda)\}$ and a global upper bound for L'_{SRMP} is given by

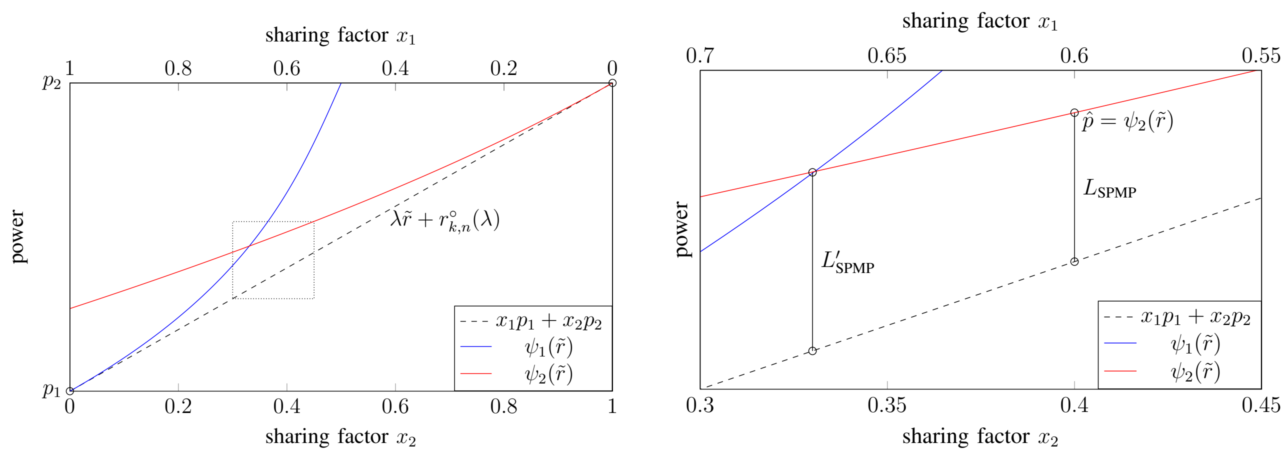
$$L''_{\text{SRMP}} = \max_n \sup_p \left\{ r_n^{\text{conv}}(p) - r_n(p) \right\}.$$

Similiarly, the power level of the convex closure of ψ_n at data rate r is given by the best tangential approximation $\psi_n^{\text{conv}}(r) = \sup_\lambda \{\lambda r + \psi^\circ_n(\lambda)\}$. It follows that

$$L''_{\text{SPMP}} = \max_n \sup_r \left\{ \psi_n(r) - \psi_n^{\text{conv}}(r) \right\}$$

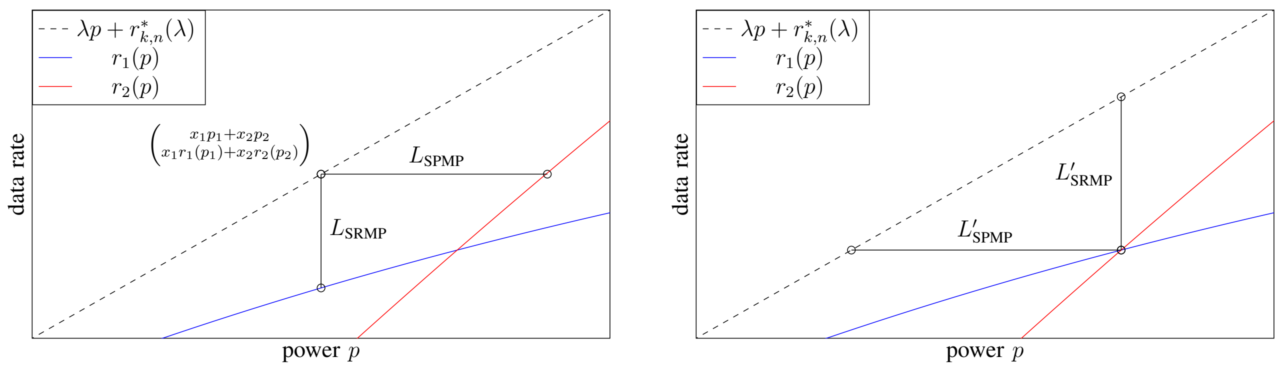
is an a-priori upper bound for L'_{SPMP}.

The terms L''_{SRMP} and L''_{SPMP} are a-priori bounds for the absolute performance loss suffered when employing a dual method and rounding. However, there is no closed-form expression for L'' and it might not be practical to compute it. Furthermore, for the purpose of a general analysis, a bound



(a) (b)

Fig. 3. Visualization of the absolute performance loss induced by rounding the SPMP. (a) The values in the interval $[p_1, p_2]$ in which power is lost due to rounding can be considered as a function of x_1 and x_2. (b) Zooming in on the point corresponding to $(\tilde{r}, \hat{p})$ for the exemplary $[x_1, x_2] = [0.6, 0.4]$ shows that L_{SPMP} depends on x_1 and x_2, whereas L'_{SPMP} is an independent upper bound.



(a) (b)

Fig. 4. Comparison of L_{SRMP}, L_{SPMP}, L'_{SRMP}, and L'_{SPMP}. (a) The rate gap L_{SRMP} is measured vertically, while the power difference L_{SPMP} is measured horizontally. (b) The worst-case bounds L'_{SRMP} and L'_{SPMP} are given by the maximum vertical and horizontal gap between $\max\{r_1, r_2\}$ and its convex closure, respectively. Both gaps are maximal at $r_1(p) = r_2(p)$ due to the concavity of the rate functions.

on the relative performance (as opposed to the absolute) is preferable. For the SRMP, the obtained rate is lower bounded by $P^{\text{abs}}_{\text{SRMP}} = r^*_{\text{SRMP-S}} - L''_{\text{SRMP}}$. It follows that the relative performance $P^{\text{rel}}_{\text{SRMP}}$ satisfies

$$P^{\text{rel}}_{\text{SRMP}} = \frac{P^{\text{abs}}_{\text{SRMP}}}{r^*_{\text{(SRMP)}}} \geq \frac{r^*_{\text{(SRMP-S)}} - L''_{\text{SRMP}}}{r^*_{\text{(SRMP-S)}}}$$
$$= 1 - \frac{L''_{\text{SRMP}}}{r^*_{\text{(SRMP-S)}}} = 1 - \frac{\rho_{\text{SRMP}}}{N}, \quad (10)$$

in which ρ_{SRMP} denotes the ratio between L''_{SRMP} and the average rate per subcarrier, $r^*_{\text{(SRMP-S)}}/N$. From a strictly formal perspective, ρ_{SRMP} can take almost arbitrary values. One can construct particularly adverse scenarios in which the affected subcarrier not only loses a large part of its data rate, but also majorly contributes to the overall objective, leading to large values of ρ_{SRMP}. However, in practice it can be assumed that a) the subcarrier is affected in a minor way, and that b) the affected subcarrier is effectively random with regards to its contribution to the overall objective. Under these assumptions, $\rho_{\text{SRMP}} \ll 1$ holds, but even a high value of $\rho_{\text{SRMP}} = 1$, which corresponds to the outage of an average subcarrier, results in $P^{\text{rel}}_{\text{SRMP}} \geq 1 - 1/N$, which is close to 100% for any practical number of subcarriers N.

For the SPMP, the power spent by applying the dual method above and rounding is upper bounded by $P^{\text{abs}}_{\text{SPMP}} = p^*_{\text{(SPMP-S)}} + L''_{\text{SPMP}}$. Because the SPMP is a minimization problem, the relative performance $P^{\text{rel}}_{\text{SPMP}}$ describes how much more power is spent compared to the optimal solution $p^*_{\text{(SPMP)}}$. The closer $P^{\text{rel}}_{\text{SPMP}} \geq 1$ is to 1, the better the performance. An upper bound for the relative performance is given by

$$P^{\text{rel}}_{\text{SPMP}} = \frac{P^{\text{abs}}_{\text{SPMP}}}{p^*_{\text{(SPMP)}}} \leq \frac{p^*_{\text{(SPMP-S)}} + L''_{\text{SPMP}}}{p^*_{\text{(SPMP-S)}}}$$
$$= 1 + \frac{L''_{\text{SPMP}}}{p^*_{\text{(SPMP-S)}}} = 1 + \frac{\rho_{\text{SPMP}}}{N}, \quad (11)$$

in which ρ_{SPMP} denotes the ratio between L''_{SPMP} and the average power $p^*_{\text{(SPMP-S)}}/N$ spent per subcarrier. This result is analogous to (10), except that $P^{\text{rel}}_{\text{SPMP}}$ is increased by spending additional power rather than decreased by losing part of the overall data rate. Under the same practical assumptions as above it again holds that $\rho_{\text{SPMP}} \ll 1$. A value of $\rho_{\text{SPMP}} = 1$ leads to $1 \leq P^{\text{rel}}_{\text{SPMP}} \leq 1 + 1/N \approx 100\%$.

The above bounds (10) and (11) show that the relative performance loss for both the SRMP and the SPMP is of the order of $1/N$, which tends to 0 as N grows. This is a sensible result as at most one subcarrier is affected by the rounding. In practice, the affected subcarrier can be expected to operate at an average power and rate and to be only affected minimally by the rounding, leading to relative performance losses far below $1/N$. Both methods are asymptotically optimal in the sense that the performance approaches that of the discrete optimum as the number of subcarriers tends to infinity, but can be expected to yield great performance for even a moderate number of subcarriers.

Relative performance in this regime make dual methods a great choice for RAPs under a global power budget (GBP) or a global rate demand (GRD). As the above performance analysis is independent of the global constraint itself, performance in the same regime can be expected from related RAPs with a single non-(MC) constraint as well. We analyze the effect of multiple non-(MC) constraints in the following section, in which individual power budgets and rate demands create a constraint set of size $K+N$.

IV. Resource Allocation under Individual Constraints

A. Sum Rate Maximization with Individual Power Budgets

We extend the results of Section III-A to the case of individual power budgets. Given power constraints $p_k^{\max} > 0$ for $k = 1, \ldots, K$, the SRMPI can be formulated as follows:

$$\text{(SRMPI)} \quad \underset{p_{k,n}\geq 0,\, x_{k,n}\in\{0,1\}}{\text{maximize}} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} r_{k,n}(p_{k,n})$$

$$\text{subject to} \quad \sum_{n=1}^{N} x_{k,n} p_{k,n} \leq p_k^{\max}, \quad k=1,\ldots,K, \qquad \text{(IPB)}$$

$$\sum_{k=1}^{K} x_{k,n} = 1, \quad n = 1,\ldots,N. \qquad \text{(MC)}$$

In the above problem, the global power budget constraint (PB) has been replaced by K individual power budgets (IPB). The SRMPI with subcarrier sharing is obtained by relaxing $x_{k,n} \in \{0,1\}$:

$$\text{(SRMPI-S)} \quad \underset{p_{k,n}\geq 0,\, x_{k,n}\geq 0}{\text{maximize}} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} r_{k,n}(p_{k,n})$$
$$\text{subject to} \quad \text{(IPB) and (MC).}$$

Proposition 5. *The dual problems of (SRMPI) and (SRMPI-S) coincide. The dual problem is*

$$\textit{(SRMPI-D)} \quad \underset{\lambda_1,\ldots,\lambda_K}{\textit{minimize}} \quad \sum_{k=1}^{K}\lambda_k p_k^{\max} + \sum_{n=1}^{N}\max_k r_{k,n}^{*}(\lambda_k)$$
$$\textit{subject to} \quad \lambda_k \geq 0, \quad k=1,\ldots,K.$$

Problems (SRMPI-S) and (SRMPI-D) have zero duality gap and a solution to (SRMPI-S) can be obtained from the dual optimum.

Proof: *Applying the transformation $q_{k,n} = x_{k,n}p_{k,n}$ to (SRMPI-S) results in an equivalent convex problem. We omit the remaining proof as it is an extension of the proof of Prop. 2.* ■

We present and analyze the SPMPI in the next subsection before jointly evaluating the performance of dual methods for both problems in Section IV-C.

B. Sum Power Minimization with Individual Rate Demands

In this section, we extend the results of Section III-B to the case of individual rate demands. Given rate demands $r_k^{\min}$ for $k = 1, \ldots, K$, the SPMPI can be formulated as follows:

$$\text{(SPMPI)} \quad \underset{p_{k,n}\geq 0, x_{k,n}\in\{0,1\}}{\text{minimize}} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} p_{k,n}$$

$$\text{subject to} \quad \sum_{n=1}^{N} x_{k,n} r_{k,n}(p_{k,n}) \geq r_k^{\min},\ k=1,\ldots,K, \quad \text{(IRD)}$$

$$\sum_{k=1}^{K} x_{k,n} = 1, \quad n=1,\ldots,N. \qquad \text{(MC)}$$

The SPMPI with subcarrier sharing is obtained by relaxing $x_{k,n} \in \{0,1\}$:

$$\text{(SPMPI-S)} \quad \underset{p_{k,n}\geq 0, x_{k,n}\geq 0}{\text{minimize}} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} p_{k,n}$$
$$\text{subject to} \quad \text{(IRD) and (MC).}$$

Proposition 6. *The dual problems of (SPMPI) and (SPMPI-S) coincide. The dual problem is*

$$\textit{(SPMPI-D)} \quad \underset{\lambda_1,\ldots,\lambda_K}{\textit{maximize}} \quad \sum_{k=1}^{K}\lambda_k r_k^{\min} + \sum_{n=1}^{N}\min_k r_{k,n}^{\circ}(\lambda_k)$$
$$\textit{subject to} \quad \lambda_k \geq 0, \quad k=1,\ldots,K.$$

Problems (SPMPI-S) and (SPMPI-D) have zero duality gap and a solution to (SPMPI-S) can be obtained from the dual optimum.

Proof: *Applying the transformation $q_{k,n} = x_{k,n}p_{k,n}$ to (SPMPI-S) yields an equivalent convex problem. We omit the proof as it is an extension of the proof of Prop. 3.* ■

C. Performance under Individual Constraints

The convexity of (SRMPI-S) and (SPMPI-S) means that a bound on the performance gap of a dual method is given by the difference in objective between a relaxed solution and a rounded feasible point of the original problem. As before, the number of nonbinary components in solutions to (SRMPI-S) and (SPMPI-S), respectively, is key. In the previous section, the number of shared subcarriers was upper bounded by 1. With K individual constraints, we obtain the following result:

Proposition 7. *Let (P-S) denote a feasible instance of either (SRMPI-S) or (SPMPI-S). Then, there exists a solution $\{x_{k,n}, p_{k,n}\}_{k,n}$ to (P-S) that satisfies*

1) $c_{\{1\}} \geq N - K$,
2) $c_{(0,1)} \leq K + N - c_{\{1\}} \leq 2K$.

In other words, there are at most K subcarriers for which $\{x_{k,n}\}_{k,n}$ has nonbinary components. Furthermore, the total number of nonbinary components does not exceed $2K$.

Proof: *Fixing power values $\{p_{k,n}\}_{k,n}$ for (P-S) results in an LP of the form of Cor. 1, which has K (IPB) or (GRD) constraints and N (MC) constraints. Thus, there exists a solution to (P-S) with $\{x_{k,n}\}_{k,n}$ such that $c_{(0,1]} \leq K+N$. It follows that*

$$2\left(N - c_{\{1\}}\right) \leq c_{(0,1)} = c_{(0,1]} - c_{\{1\}} \leq K + N - c_{\{1\}} \qquad (12)$$

and subtracting $N-c_{\{1\}}$ from both sides yields $N-c_{\{1\}} \leq K$, which is equivalent to 1) and yields 2) from (12). ■

Just as Prop. 4 in the previous section, Prop. 7 shows that there is a direct correspondence between the number of non-(MC) constraints and the number of shared subcarriers in a solution to (P-S). Introducing subcarrier sharing to one of the RAPs with K individual constraints results in a relaxed solution in which at most K subcarriers are shared between users.

As before, the goal of this section is to show that the relative performance loss is of the order of K/N because rounding occurs on K out of N subcarriers. In practice, the performance losses on each subcarrier can be expected to be relatively small compared to the overall rate obtained/power spent on each subcarrier, leading to even better relative performance.

In the previous section, the constraints of (SRMP) and (SPMP) were interchangeable. However, as it turns out, there is a major difference between the individual power budgets (IPB) and the individual rate demands (IRD). While each user can always save power to satisfy (IPB), there might not be sufficient resources to jointly satisfy (IRD). This means that from a formal perspective, feasibility and thus performance of dual methods for the (SPMPI) can not be guaranteed. We expand on this negative result before analyzing the performance of the (SRMPI) in detail.

Let $\{p_{k,n}, x_{k,n}\}$ be a solution to (SPMPI-S) satisfying the bounds of Prop. 7. As before, a binary set $\{x_{k,n}\}_{k,n}$ means that this solution is also optimal for (SPMPI). Otherwise, there are up to K shared subcarriers and up to 2^K possible subcarrier assignments obtainable through rounding. In practice, most or all of these assignments should yield feasible results once the well-known waterfilling procedure is applied to the final allocation.

However, the nature of the individual rate demands (IRD) of the SPMPI means that there can be subcarrier assignments which make it impossible to guarantee the existence of a feasible rounding. In theory, roundings can lead to single users being left without a sufficient number of subcarriers to satisfy their rate demands $r_k^{\min}$, $k = 1, \ldots, K$.

This is because allocating a shared subcarrier to only one user necessarily reduces the data rate of at least one other user. As the optimal solution to (SPMPI-S) satisfies all rate demands with equality, this makes the rounded solution infeasible for this user. There are two ways to compensate this effect: The first way is to allocate another shared subcarrier to the user, but that potentially causes another user to drop below their rate demand. Depending on the total number of subcarriers available, the number of users and subcarriers shared, it might not be possible to satisfy all rate demands this way.

The second way to compensate for lost data rate of a user is to increase the power spent on their remaining subcarriers. However, even under the assumption that there is a sufficient number of subcarriers for each user, the presented approach offers no way to bound the amount of power that is required to adjust for the rate loss introduced through rounding. This is why the existence of a feasible rounding for the SPMPI can not be formally guaranteed.

In practice, the rate decrease from losing a previously shared subcarrier is counterbalanced by two factors. First, the user might be assigned additional subcarriers. Second, a small increase in power on each subcarrier according to the waterfilling procedure yields the most power-efficient way to generate data rate, even if slightly more expensive than the optimal shared solution to (SPMPI-S). There might, however, be severe differences in performance between the available roundings, which again shows the danger of interpreting the results of a dual method as a unique solution to the primal problem.

It still holds that an SPMPI system with full subcarrier sharing converges to a system with unique subcarrier usage in the sense that the ratio of shared subcarriers is K/N, which tends to 0 as N grows. As the error introduced by rounding only negatively affects the shared subcarriers, the relative error is of the order of K/N as well, with the actual value being much lower as the subcarriers are still being utilized after reallocation. Under the assumption that efficiently finding a feasible rounding is always possible in a practical system with a sufficient number of subcarriers, dual methods for the SPMPI are as performant and asymptotically optimal as those for the SRMPI, which are covered next.

Let $\{p_{k,n}, x_{k,n}\}_{k,n}$ be a solution to (SRMPI-S) satisfying the bounds of Prop. 7. In the case that $\{x_{k,n}\}_{k,n}$ is binary, this solution is optimal for (SRMPI) because (SRMPI-S) relaxes (SRMPI). Otherwise, the nonbinary components can be rounded to obtain feasible points for (SRMPI). For each shared subcarrier n, let j denote one of the users with $x_{j,n} \in (0,1)$. Then, a feasible rounding is given by

$$(x_{k,n}, p_{k,n}) \mapsto \begin{cases} (1, x_{k,n} p_{k,n}), & \text{if } k = j, \\ (0,0), & \text{else.} \end{cases}$$

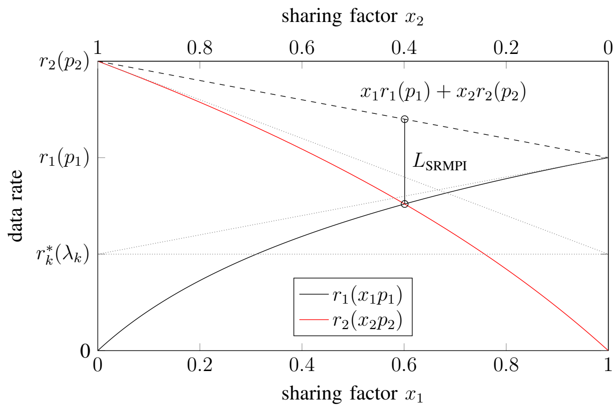
For user j, the amount of power spent on subcarrier n remains the same. All the other users do not spend any power on subcarrier n, which means that the (IPB) constraints are satisfied. Different from Section III-A, the rounding above does not make full use of all the power budgets $p_k^{\max}$ for each user. However, it shows that all assignments of previously shared subcarriers to one user each are feasible. Once an assignment is fixed, optimal operating points $\{p_{k,n}\}_{k,n}$ can be found by the well-known waterfilling procedure.

To bound the performance, we start by analyzing a single shared subcarrier n. To simplify notation, we drop the subcarrier index n in the following power-rate formulas. Let $\mathcal{J}$ denote the set of users sharing the subcarrier. An upper bound for the performance loss is given by assigning the subcarrier to the user with the highest data rate. Define

$$\mathcal{F} := \left\{ \sum_{k \in \mathcal{J}} x_k r_k(p_k) - \max_{k \in \mathcal{J}} \{r_k(x_k p_k)\} \,\middle|\, \sum_{k \in \mathcal{J}} x_k = 1 \right\}.$$

Then, the upper bound is given by $L_{\text{SRMPI}} = \sup \mathcal{F}$, which describes the maximum difference between the convex combination of the rate values $r_k(p_k)$ and the rate obtainable by a single user. Note that $x_k r_k(p_k) \leq r_k(x_k p_k)$ holds due to the concavity of the rate functions. Define

$$\mathcal{G} := \left\{ \sum_{k \in \mathcal{J}} x_k r_k(p_k) - \max_{k \in \mathcal{J}} \{x_k r_k(p_k)\} \,\middle|\, \sum_{k \in \mathcal{J}} x_k = 1 \right\}.$$



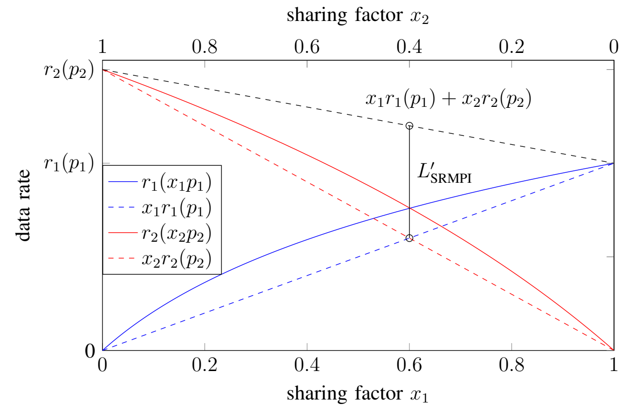


Fig. 5. Visualization of the bounds on the absolute performance loss induced by rounding the SRMPI in the case that two users share a subcarrier. (a) Upper bound $L_{\text{SRMPI}} = \sup \mathcal{F}$. (b) Simplified upper bound $L'_{\text{SRMPI}} = \sup \mathcal{G}$.

Then, a weaker but simpler bound is given by $L'_{\text{SRMPI}} = \sup \mathcal{G}$. Fig. 5 shows the above bounds for $\mathcal{J} = \{1, 2\}$. The advantage of L'_{SRMPI} is that finding the supremum is equivalent to solving an LP. Evaluating the Karush-Kuhn-Tucker conditions of this LP shows that the optimum is achieved when $x_k r_k(p_k) = x_j r_j(p_j)$ for all k and j, which translates to

$$x_k = \frac{1}{r_k(p_k)} \cdot \frac{1}{\sum_{j\in\mathcal{J}} \frac{1}{r_j(p_j)}}$$

for all $k \in \mathcal{J}$ and

$$L'_{\text{SRMPI}} = \frac{|\mathcal{J}| - 1}{\sum_{j=1}^{J} \frac{1}{r_j(p_j)}}.$$

Let $J := |\mathcal{J}|$. Then, the above term can be bounded by

$$L'_{\text{SRMPI}} \leq \frac{J-1}{J \cdot \frac{1}{\max_{k\in\mathcal{J}}\{r_k(p_k)\}}} = \frac{J-1}{J} \cdot \max_{k\in\mathcal{J}} \{r_k(p_k)\},$$

which will provide the basis for the following performance analysis. Note that $(J-1)/J$ is monotonically increasing and takes values between $1/2$ and $(K-1)/K$. Thus, the performance guarantee is better the fewer users share a single subcarrier. However, there is a trade-off as the number of subcarriers being shared and the number of users sharing them is related as shown in Prop. 7.

Let S denote the number of shared subcarriers and assume that the subcarriers are ordered such that $n = 1, \ldots, S$ are the shared subcarriers. According to Prop. 7, the amount of shared subcarriers $S = N - c_{\{1\}}$ satisfies $S \leq K$. Furthermore, the number of users sharing subcarriers is upper bounded by $K + S$. Let $\mathcal{J}_n$ denote the set of users sharing subcarrier $n = 1, \ldots, S$. With $J_n := |\mathcal{J}_n|$, the performance loss sums up to

$$\sum_{n=1}^{S} \frac{J_n - 1}{J_n} \max_{k\in\mathcal{J}_n} \{r_{k,n}(p_{k,n})\} \leq \max_{1\leq n\leq S,\, k\in\mathcal{J}_n} \{r_{k,n}(p_{k,n})\} \sum_{n=1}^{S} \frac{J_n - 1}{J_n}.$$

For a worst-case analysis, the term $\sum_{n=1}^{S}(J_n - 1)/J_n$ has to be maximized under the condition $\sum_{n=1}^{S} J_n = S + K$. The optimal choice is to distribute users between shared subcarriers as evenly as possible. Setting $J_n = (S+K)/S$ is optimal if J_n takes integer values, but also yields an upper bound for the problem for non-integer J_n. Thus, it holds that

$$\sum_{n=1}^{S} \frac{J_n - 1}{J_n} \leq S \cdot \frac{\frac{K}{S}}{\frac{S+K}{S}} = \frac{SK}{S+K} \leq \frac{K^2}{2K} = \frac{K}{2}.$$

It follows that the performance loss for the overall problem is bounded by

$$L''_{\text{SRMPI}} = \frac{K}{2} \cdot \max_{1\leq n\leq S,\, k\in\mathcal{J}_n} \{r_{k,n}(p_{k,n})\}, \tag{13}$$

which is also obtained when analyzing the special case of $S = K$ and $J_n = 2$ for $n = 1, \ldots, S$. Note that the maximization in (13) is performed over the shared subcarriers and all users sharing those subcarriers, with $p_{k,n}$ given by the output of the dual method. However, it also serves as an a-priori bound in the sense that the absolute performance loss is of the order of a single subcarrier usage times $K/2$. With similar arguments as in Section III-A, this allows to bound the relative performance by

$$P^{\text{rel}}_{\text{SRMPI}} \geq \frac{r^*_{\text{(SRMPI-S)}} - L''_{\text{SRMPI}}}{r^*_{\text{(SRMPI-S)}}} = 1 - \frac{L''_{\text{SRMPI}}}{r^*_{\text{(SRMPI-S)}}} = 1 - \frac{\rho_{\text{SRMPI}} \cdot K}{2N},$$

in which ρ_{SRMPI} denotes the ratio between $\max_{1\leq n\leq S,\, k\in\mathcal{J}_n} r_{k,n}(p_{k,n})$ and the average rate per subcarrier, $r^*_{\text{(SRMPI-S)}}/N$. As before, ρ_{SRMPI} can take almost arbitrary values in theory, but it is reasonable to assume $\rho_{\text{SRMPI}} \leq 2$ to hold in practice. It then follows that $P^{\text{rel}}_{\text{SRMPI}} \geq 1 - K/N$, which is a weaker bound as for the SRMP for two reasons. Primarily, the SRMPI has up to K subcarriers being shared as opposed to one in the SRMP case. Furthermore, the bound L''_{SRMPI} is relatively weak due to combining multiple worst-case bounds. However, it still holds that $P^{\text{rel}}_{\text{SRMPI}}$ tends towards 100% as N grows. Furthermore, the number of users K in a practical system can be expected to be significantly lower than the number of subcarriers. That said, solving the SRMPI with dual methods can still be expected to yield relative performance close to 100% for systems with a moderate number of subcarriers and is asymptotically optimal as the number of subcarriers tends to infinity.

V. Conclusion

This work shows that applying dual methods is equivalent to solving a linearly-relaxed, "timeshared" formulation of the RAP. It is shown that the relaxed RAPs are convex. This result constitutes a missing piece in the search for optimality arguments regarding dual methods. An immediate consequence of the above is that the duality gap encountered with dual methods is equal to the error introduced when rounding a timeshared solution towards a solution with unique subcarrier usage, making it feasible for the original problem. This work shows that the number of subcarriers being shared in a fully timeshared solution is bounded by the number of constraints governing (individual or global) power budgets and rate demands.

This result is then applied to obtain feasible roundings based on the output of dual methods. Further analysis of the problem leads to absolute and relative performance guarantees. Under mild conditions regarding balanced, nondegenerate subcarrier usage, the relative error is formally bounded by $1/N$ and K/N, respectively. This paper covers four RAPs: The SRMP with a global power budget, the SRMP with individual power budgets, the SPMP with a global rate demand, and the SRMP with individual rate demands. By formally combining results from several areas of optimization theory, it serves as a blueprint for obtaining similar results for related problems in resource allocation and beyond.

References

[1] J. Campello, "Optimal discrete bit loading for multicarrier modulation systems," in *Proc. 1998 IEEE International Symp. Inf. Theory*, p. 193.
[2] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, July 2006.
[3] J. P. Aubin and I. Ekeland, "Estimates of the duality gap in nonconvex optimization," *Mathematics Operations Research*, vol. 1, no. 3, pp. 225–245, 1976.
[4] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*. Athena Scientific, 1996.
[5] K. Seong, M. Mohseni, and J. Cioffi, "Optimal resource allocation for OFDMA downlink systems," in *Proc. 2006 IEEE International Symp. Inf. Theory*, pp. 1394–1398.
[6] J. Huang, V. Subramanian, R. Agrawal, and R. Berry, "Downlink scheduling and resource allocation for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 288–296, Jan. 2009.
[7] C. Liu, A. Schmeink, and R. Mathar, "Dual optimal resource allocation for heterogeneous transmission in OFDMA systems," in *2009 IEEE Globecom.*
[8] X. Wang and G. Giannakis, "Resource allocation for wireless multiuser OFDM networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4359–4372, July 2011.
[9] I. Wong and B. Evans, "Optimal downlink OFDMA resource allocation with linear complexity to maximize ergodic rates," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 962–971, Mar. 2008.
[10] ——, *Resource Allocation in Multiuser Multicarrier Wireless Systems*. Springer, 2008.
[11] ——, "Optimal resource allocation in the OFDMA downlink with imperfect channel knowledge," *IEEE Trans. Commun.*, vol. 57, no. 1, pp. 232–241, Jan. 2009.
[12] H. Zhu and J. Wang, "Chunk-based resource allocation in OFDMA systems–part II: joint chunk, power and bit allocation," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 499–509, Feb. 2012.
[13] S. Görtzen and A. Schmeink, "Optimality of dual methods for discrete multiuser multicarrier resource allocation problems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3810–3817, Oct. 2012.
[14] T. Cover and J. Thomas, *Elements of Information Theory*, ser. Wiley Series in Telecommunications and Signal Processing. Wiley-Interscience, 2006.
[15] S. T. Chung and A. Goldsmith, "Degrees of freedom in adaptive modulation: a unified view," *IEEE TRans. Commun.*, vol. 49, no. 9, pp. 1561–1571, Sept. 2001.
[16] H. Kellerer, U. Pferschy, and D. Pisinger, *Knapsack Problems*. Springer, 2004.
[17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
[18] J.-B. Hiriart-Urruty and C. Lemaréchal, *Fundamentals of Convex Analysis*. Springer Verlag, Heidelberg, 2001.
[19] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems*, ser. Studies in Mathematics and Its Applications. Elsevier Science, 1976.
[20] G. L. Nemhauser and L. A. Wolsey, *Integer and Combinatorial Optimization*. Wiley-Interscience, 1988.

Simon Görtzen received his Diploma degree in mathematics with a minor in computer science from RWTH Aachen University, Germany, in 2008. He is a Ph.D. student at the UMIC Research Centre at RWTH Aachen University, pursuing a degree in electrical engineering and information technology. His research interests are OFDMA systems and resource allocation problems, convex optimization techniques and algorithm design.

Anke Schmeink received the Diploma degree in mathematics with a minor in medicine and the Ph.D. degree in electrical engineering and information technology from RWTH Aachen University, Germany, in '02 and '06, respectively. She worked as a research scientist for Philips Research before joining RWTH Aachen University in '08 where she is an associate professor since '12. She spent several research visits with the University of Melbourne, and with the University of York. Anke Schmeink is a member of the Young Academy at the North Rhine-Westphalia Academy of Science. Her research interests are in information theory, systematic design of communication systems and bioinspired signal processing.